

Associative Concept Learning, Stimulus Equivalence, and Relational Frame Theory:
Working out the Similarities and Differences between Human and Non-Human Behavior

Sean Hughes and Dermot Barnes-Holmes

National University of Ireland Maynooth, Ireland

Author Note

Sean Hughes and Dermot Barnes-Holmes, Department of Psychology, National University of Ireland, Maynooth. Please address correspondence to Sean Hughes or Dermot Barnes-Holmes, National University of Ireland Maynooth, County Kildare, Ireland. E-mail: sean.hughes@nuim.ie or dermot.barnes-holmes@nuim.ie. This paper was funded by a Government of Ireland fellowship from the Irish Research Council to the first author.

Abstract

In their review article Zentall et al. propose that nonhumans can come to relate stimuli based on their physical properties (*perceptual concept learning*) or the relationship established between or among physically related stimuli (*relational concept learning*). At the same time, they draw upon findings from within the animal learning literature in order to argue that nonhumans can also derive untrained yet predictable relations between stimuli in the absence of direct training (*associative concept learning*). We are both intrigued and excited by the body of work contained in this paper and believe that it may accelerate our understanding of animal as well as human cognition in several ways. Nevertheless, a number of important questions still need to be addressed before we can conclude that associative concept learning in nonhumans is functionally similar to that observed in our own species.

Associative Concept Learning, Stimulus Equivalence, and Relational Frame Theory:

Working out the Similarities and Differences Between Human and Non-Human Behavior

The question of what makes humans unique has attracted considerable attention within the behavioral sciences. Throughout much of the past century it was assumed that those learning principles identified in non-humans could stretch to, and account for, much of complex human behavior (see Dymond, Roche, & Barnes-Holmes, 2003; Hayes, 1987; Hayes, Barnes-Holmes & Roche, 2001). This ‘continuity assumption’ served as an ‘intellectual rudder’ guiding early work in the field, with researchers focusing pragmatically on non-humans in order to identify general learning principles that could predict-and-influence the actions of our own species. In many respects, this analytic strategy was a successful one, yielding concepts that appear to apply equally to humans and non-humans alike (e.g., reinforcement, punishment, generalization, discrimination, extinction, recovery and habituation). The continuity assumption seemed to hold true.

However, when researchers turned their attention to those hallmarks of human language (and cognition) a number of important findings started to emerge, findings that hinted at learning processes or principles that may be unique to, or largely elaborated in, some species relative to others. For instance, verbal behavior stubbornly refused to be analyzed in direct contingency terms and early attempts to do so (e.g., Skinner, 1957) did not yield a progressive program of research. Furthermore, a growing body of work on instructional control (or rule-governed behavior) and stimulus equivalence revealed that humans consistently respond in ways that differ from their non-humans counterparts (Galizio, 1979; Sidman, 1971). Thus it seemed as if - at least in some cases - the continuity assumption was beginning to crack and strain.

Fast forward forty years and these species-related differences are still the subject of intense controversy and debate (e.g., Lionello-DeNolf, 2009; Kastak, Schusterman & Kastak,

2001; Hayes, 1989; Sidman et al., 1982; Urcuioli, 2008). Some researchers have argued that the behavioral processes involved in human language are functionally similar to those that are involved in human responding on equivalence tasks (Hayes et al., 2001). The key behavioral process (known as arbitrarily applicable relational responding; AARR) refers to the capacity to relate stimuli in ways that (a) do not depend on the formal properties of the to-be-related stimuli and (b) that are controlled by aspects of the context that have been abstracted so that they can be arbitrarily applied in a wide variety of ways. For instance, if humans learn that *A-Same-B* and that *B-Same-C*, they will form a number of bi-directional relations between these stimuli in the absence of any training or instruction to do so. That is, they will act as if *B-Same-A* and *C-Same-B* (mutual entailment) as well as if *A-Same-C* and *C-Same-A* (combinatorial entailment). Furthermore, the (psychological) properties of those stimuli will also change in-line with the manner in which they were related (known as a transformation of function). Thus if an aversive function is established for A by repeatedly pairing it with an electric shock, B and C will also acquire aversive properties despite the fact that they were never paired with shocks at any point in time (Dougher, Augustson, Markham, Greenway, & Wulfert, 1994).

Although humans appear to learn AARR early on in their development, existing evidence for mutual and combinatorial entailment has been difficult to find in other species such as pigeons (Lionello-DeNolf, & Urcuioli, 2002), chimpanzees, and baboons (Dugdale & Lowe, 2000, Hayes, 1989; Sidman et al., 1982). Indeed, studies demonstrate that even after extensive training non-humans find it difficult to demonstrate the simplest form of AARR (i.e., symmetry or mutual entailment). Moreover, in those cases where positive evidence has been obtained for symmetry responding, test performances can be explained in ways that (a) do not involve AARR, (b) are only present in a small sub-section of the sample, or (c) are emitted with unacceptably low levels of accuracy (Lionello-DeNolf, 2009). Nevertheless,

more recent papers (including the target article) argue that these failures to observe AARR in animals stem from properties of the test procedures themselves and non-humans are capable of AARR under a set of highly specific conditions.

The basic argument is that non-humans cannot only relate stimuli based on their physical properties (*perceptual concept learning*) or as a function of the relationship established between or among physically related stimuli (*relational concept learning*), but also derive relations that were never directly trained or instructed in the past (*associative concept learning*). In this latter case, the organism is argued to relate stimuli in ways that do not depend on their physical properties and in such a way that these stimuli come to be functionally interchangeable with one another (i.e., respond in an arbitrarily applicable fashion). These performances are argued to mirror those observed in humans - and by implication - reflect a learning process that stretches across the species divide. This claim is backed by recent findings from within the animal learning literature on acquired equivalence and symmetry responding, in which nonhumans purportedly relate stimuli in untrained yet predictable ways.

Overall, we are both intrigued and excited by the body of work contained in this review and believe that it may accelerate our understanding of animal as well as human cognition in several ways. Before continuing, however, it is important to recognize that these empirical developments are not incompatible or inconsistent with Relational Frame Theory (RFT). Actually it's quite the opposite. Almost 20 years ago, Barnes and Roche (1996) wrote:

RFT recognizes that very limited forms of derived behavior may occur without a history of explicitly reinforced equivalence responding (see Hayes & Wilson, 1993; see also Zentall and Urcuoli, 1993, for evidence that suggests derived behavior may occur in nonhumans). Nevertheless, these types of behavior are normally defined as largely respondent because they do not emerge from an appropriate history of arbitrarily applicable relational responding (note, this is a purely functional distinction). Respondent behavior is, however, considered to be an important foundation for relational framing (see Barnes, 1994, for a detailed discussion of this issue in terms of indirect reflexivity) (p. 501).

It is also interesting to note, that in that same article Barnes and Roche presented some speculative RFT-based analyses concerning the relationship between identity matching (or reflexivity) and derived symmetry responding, concluding that the “issue will clearly require some very precise and delicate analyses by those researchers who specialize in nonhuman stimulus control” (p. 502). The empirical work that is reviewed in the current target article provides, in our view, some of the finest examples of the very research that Barnes and Roche called for back in the mid-1990s.

Despite these positives, we believe that an argument for AARR in non-humans currently faces a number of challenges. We know that AARR is characterized (at least in humans) by its flexibility. For instance, humans can respond to stimuli as if they are equivalent or symmetrically related based on verbal instructions (Smeets, Dymond & Barnes-Holmes, 2000), conditional discrimination tasks (Dougher et al., 1994) and respondent-like training procedures (Leader & Barnes-Holmes, 2001). Equivalence (and other derived relations) also emerge regardless of whether a Many-to-One (MTO), One-To-Many (OTM) or linear training and testing design is employed (i.e., where A is related to B and B related to C; Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Holth, 1997). No strict temporal or spatial ordering between the sample and comparison stimuli must be implemented nor is concurrent identity or successive matching training a prerequisite in order to observe these effects in verbal humans.

If Zentall et al. are correct, and non-humans are capable of AARR in the same way as their human counterparts, then they appear to be restricted to the most rudimentary features of that behavior, which emerge only under the strictest of experimental conditions. The work highlighted in their review suggests that symmetry may depend on successive matching in which the sample and comparison stimuli are presented one after the other in the same location (and close in time). This seems to differ functionally from the symmetry

performances observed in humans that emerge regardless of the training protocol used. Likewise, the difficulty of observing equivalence performances on OTM (and perhaps even linear training and testing designs) also differs from comparable performances in humans (Arntzen et al., 2010). When taken together, the reviewed research does not (convincingly) demonstrate the flexibility that is characteristic of AARR in our own species.

It also seems important to acknowledge that humans are not shackled to associative concept learning (or relating stimuli on the basis of equivalence) in their interactions with the environment but can instead behave as if stimuli are related to one another in many different ways. Indeed, findings indicate that this ability to respond in an arbitrarily applicable fashion enables humans to relate stimuli as opposite (Dymond, Roche, Forsyth, Whelan & Rhoden, 2008), hierarchically (Gil, Luciano, Ruiz & Valdivia-Salas, 2012), comparatively (Vitale, Barnes-Holmes, Barnes-Holmes & Campbell, 2008), deictically (McHugh & Stewart, 2012), temporally and/or causally (O'Hora et al., 2008) related (see Dymond & Roche, 2013, for a recent review). This work also suggests that when non-equivalence relations are involved, stimulus functions are not simply transferred but rather *transformed* through those relations in non-equivalent ways. For instance, if humans learn that *A-Opposite-B-Same-C* and *A-Opposite-Shock* they may come to approach A and avoid B or C (Whelan & Barnes-Holmes, 2004).

Thus, while equivalence appears to be one of the first types of derived relations to emerge in a human infant's repertoire (Luciano, Gomez-Becerra, & Rodriguez-Valverde, 2007), other relations play a defining role in how we adapt to the world around us. Deictic relations ("*I-You/ Here-There/ Now-Then*") seem to dominate many self and perspective taking behaviors (McHugh & Stewart, 2012) while temporality and causality may be central to many rule-governed behaviors (Torneke et al., 2008). Even classification or categorization itself can involve stimuli that are hierarchically related in non-equivalent ways. For example,

within the superordinate category ‘foods’ there may be classes of stimuli that are related on the basis of opposition (‘poisonous vs. non-poisonous foods’), comparison (‘cheap vs. expensive foods’), deictics (‘foods I like vs. don’t like’) or difference (‘meat vs. vegetables). Developing an account of the relational responses involved in these types of complex categorization abilities, in terms of associative concept learning (equivalence) alone, will certainly be very challenging (see Dymond & Barnes, 1995, for an empirical example of AARR responding in accordance with just two relational frames -- *same* and *comparison* -- which proved difficult to explain in terms of equivalence responding).

Despite these challenges, however, we must emphasize again that we are genuinely excited and intrigued by the work presented in the target article. Although a number of important questions still need to be addressed before we can conclude AARR in nonhumans is functionally similar to that observed in our own species, we nevertheless believe that the research outlined here represents a golden opportunity to develop much needed dialogue between animal and human learning researchers on issues that are central to both traditions. Importantly, this dialogue is not a one-way street: the lessons learned in human research may stimulate developments in the animal literature and vice-versa. For instance, research on animal cognition may provide new insights into human cognition by identifying how basic forms of (respondent) learning are selected, modified, strengthened and elaborated by the (verbal) operant contingencies that RFT argues are crucial for the emergence of increasingly complex forms of AARR. Animal preparations and populations may also allow us to examine questions about AARR that cannot be tackled with humans for ethical and practical reasons. In particular, the research considered in the target article could help us understand how young infants might transition from very basic to more complex forms of AARR.

Recent work on human cognition could also unlock new insights into animal learning. It may well be that there is some level of relational complexity, contextual control or

generalizability that humans are capable of that is not seen elsewhere in the animal kingdom. For instance, to what degree are other species restricted to equivalence (or associative concept) learning and can they respond to stimuli as opposite, more than/less than, hierarchically, temporally or casually related in an arbitrarily applicable fashion? Can they relate entire equivalence relations to other equivalence relations and form increasingly complex networks of stimulus relations (e.g., Barnes, Hegarty, & Smeets, 1997)? If so, then what implications does this have for our understanding of animal cognition, reasoning and creativity? What type, amount and order of training are required before relational responding becomes abstracted and generalizes to novel stimuli? Is it the case that this advanced type of relational learning stretches across many different evolutionary branches or is it unique to a small number of species? Again, if so, then why? Are there certain environmental or evolutionary conditions necessary in order to observe the emergence of complex forms of AARR (e.g., Hayes & Long, 2013; Smet & Byrne, in press)? Chasing these (and related) issues will not only lead to a better understanding of human and animal cognition but also identify important lines of fracture between and among species. The work presented by Zentall et al. already draws attention to potential lines of fracture between human and non-human behavior in terms of the restricted set of conditions under which non-human equivalence responding is seen to emerge. We look forward to increased collaboration and communication between animal learning and RFT researchers as we continue to explore the commonalities that bind, and the differences that separate, humans from other species in the animal kingdom.

References

- Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and the subsequent performance on the tests for stimulus equivalence. *The Psychological Record*, 60, 437–462.
- Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record*, 47, 309-320.
- Barnes, D., Hegarty, N., & Smeets, P. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior*, 14, 1-27.
- Barnes, D., & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders' commentary. *The Psychological Record*, 46, 489-507.
- Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, 62, 331-351.
- Dugdale, N., & Lowe, C.F. (2000). Testing for symmetry in the conditional discriminations of language-trained chimpanzees. *Journal of the Experimental Analysis of Behavior*, 73, 5-22.
- Dymond, S., & Roche, B. (2013). *Advances in Relational Frame Theory: Research and Application*. Oakland, CA: New Harbinger Publications.
- Dymond, S., Roche, B., & Barnes-Holmes, D. (2003). The continuity strategy, human behavior, and behavior analysis. *The Psychological Record*, 53(3), 333-347.
- Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with same and opposite relational frames. *The Psychological Record*, 58, 269-286.

- Galizio, M. (1979). Contingency-shaped and rule-governed behavior: Instructional control of human loss avoidance. *Journal of the Experimental Analysis of Behavior*, 31, 53-70.
- Gil, E., Luciano, C., Ruiz, F. J., & Valdivia-Salas, V. (2012). A Preliminary Demonstration of Transformation of Functions through Hierarchical Relations. *International Journal of Psychology and Psychological Therapy*, 12, 1-19.
- Hayes, S. C. (1987). Language and the incompatibility of evolutionary and psychological continuity. *Behavior Analysis*, 22, 49-54.
- Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 51, 385-392.
- Hayes, S.C. (2013, October 09). Variation and Selection in Rules and Rule-Governance: The Perspective of Behavioral Psychotherapy. *Social Evolution Forum*. Retrieved from <http://socialevolutionforum.com/2012/04/06/steven-c-hayes-variation-and-selection-in-rules-and-rule-governance-the-perspective-of-behavioral-psychotherapy/>.
- Hayes, S.C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum Press.
- Hayes, S.C., & Long, D. (2013). Contextual Behavioral Science, Evolution, and Scientific Epistemology. In Dymond, S & Roche, B. (2013). *Advances in Relational Frame Theory: Research & Application*. Oakland, CA: New Harbinger Publications.
- Hayes, S. C., & Wilson, K. G. (1993). Some applied implications of a contemporary behavior-analytic account of verbal events. *The Behavior Analyst*, 16, 283-301.
- Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior*, 76, 131-158.

- Leader, G., & Barnes-Holmes, D. (2001). Matching-to-sample and respondent-type training as methods for producing equivalence relations: Isolating the critical variable. *The Psychological Record*, 51, 429-444.
- Lionello-DeNolf, K.M. (2009). The search for symmetry: 25 years in review. *Learning & Behavior*, 37, 188-203.
- Lionello-DeNolf, K.M., & Urcuioli, P.J. (2002). Stimulus control topographies and tests of symmetry in pigeons. *Journal of the Experimental Analysis of Behavior*, 78, 467-495.
- Luciano, C., Gómez-Becerra, I., & Rodríguez-Valverde, M. (2007). The Role of Multiple-Exemplar Training and Naming in Establishing Derived Equivalence in an Infant. *Journal of Experimental Analysis of Behavior*, 87, 349-365.
- McHugh, L., & Stewart, I. (2012). *The self and perspective taking: Contributions and applications from modern behavioral science*. Oakland: New Harbinger Publications.
- O'Hora, D., Pelaez, M., Barnes-Holmes, D., Rae, G., Robinson, K., & Chaudhary, T. (2008). Temporal relations and intelligence: Correlating relational performance with performance on the WAIS-III. *The Psychological Record*, 58, 569-584.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14, 5-13.
- Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior*, 37, 23-44.
- Smeets, P.M., Dymond, S., & Barnes-Holmes, D. (2000). Instructions, stimulus equivalence, and stimulus sorting: Effects of sequential testing arrangements and a default option. *The Psychological Record*, 50, 339-354.

- Smet, A.F., & Byrne, R.W. (in press). African Elephants Can Use Human Pointing Cues to Find Hidden Food. *Current Biology*, <http://dx.doi.org/10.1016/j.cub.2013.08.037>.
- Torneke, N., Luciano, C., & Valdivia Salas, S. (2008). Rule-Governed Behavior and Psychological Problems. *International Journal of Psychology and Psychological Therapy*, 8(2), 141-156.
- Urcuioli, P.J. (2008). Associative symmetry, “anti-symmetry”, and a theory of pigeons' equivalence-class formation. *Journal of the Experimental Analysis of Behavior*, 90, 257–282.
- Vitale, A., Barnes-Holmes, Y., Barnes-Holmes, D., & Campbell, C. (2008). Facilitating responding in accordance with the relational frame of comparison: Systematic empirical analyses. *The Psychological Record*, 58, 365-390.
- Whelan, R., & Barnes-Holmes, D. (2004). The transformation of consequential functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior*, 82, 177-195.
- Zentall, T.R., & Urcuioli, P.J. (1993). Emergent relations in the formation of stimulus classes in pigeons. *The Psychological Record*, 43, 795-810.